Forging Emotions: A Deep Learning Experiment on Emotions and Art

Amalia Foka
School of Fine Arts, University of Ioannina
PO Box 1186, Ioannina, 45110, Greece
afoka@uoi.gr

Abstract

Affective computing is an interdisciplinary field that studies computational methods that relate to or influence emotion. These methods have been applied to interactive media artworks, but they have been focused on affect detection rather than affect generation. For affect generation, computationally creative methods need to be explored that lately have been driven with the use of Generative Adversarial Networks (GANs), a deep learning method. The experiment presented in this paper, Forging Emotions, explores the use of visual emotion datasets and the working processes of GANs for visual affect generation, i.e., to generate images that can convey or trigger specified emotions. This experiment concludes that the methodology used so far by computer science researchers to build image datasets for describing high-level concepts such as emotions is insufficient and proposes utilizing emotional networks of associations according to psychology research. Forging Emotions also concludes that to generate affect visually, merely corresponding to basic psychology findings, e.g., bright or dark colors, does not seem adequate. Therefore, research efforts should be targeted towards understanding the structure of trained GANs and compositional GANs in order to produce genuinely novel compositions that can convey or trigger emotions through the subject matter of generated images.

Keywords

Deep Learning, Affective Computing, Visual Emotion Datasets, Generative Adversarial Network (GAN).

Introduction

The notion of emotion, as we understand it today, was established through scientific research in various disciplines. Emotions have been studied from the perspective of various disciplines, including psychology, neuroscience, sociology, medicine, history as well as computer science. As a result, the question "*What is an emotion?*" will rarely generate the same answer from different scientists or individuals [1]. This is primarily due to the different basis of the emotional phenomena or theoretical issues studied from every discipline. The main questions asked for emotions are: where they originate, what physiological, behavioral, and cognitive changes they produce, what are the different expressions (e.g., facial) they induce, and what are the subjectively experienced feelings.

Emotions also have a long-standing relationship with art. Baroque art intended to appeal to the senses and emotions. Romanticism rejected the formality of Neoclassicism and, instead, embraced emotion. In Expressionism, artists depicted the world with distorted images to express their own feelings. Many theories of the expression and experience of emotion in art exist [2].

Katja Kwastek [3] has studied the aesthetic experience of the recipient in the case of interactive media artworks, which too often includes the arousal of emotions. In non-interactive works, the recipient's emotional experience is staged in the past tense; either an artist represents the emotions he/she has experienced or plans the desired emotional experience for the recipient, albeit it is mostly subjective. In contrast, interactive media artworks are located in the present tense. However, as Kwastek notes, even if the interaction process leaves scope for the unexpected, the orchestration of emotional experiences is designed in an even more explicit manner since feedback processes have to be programmed into the technical system implementing the artwork.

Consequently, affecting computing methods become extremely pertinent for orchestrating emotional experiences in interactive media artworks. However, affective computing research is, to a large extent, concentrated on detecting rather than on generating or triggering emotions. Accordingly, interactive media artworks that consider affective computing methods also mainly utilize them for affect detection.

If the problem of affect generation is considered, then computational creativity methods need to be explored, which most commonly involve Generative Adversarial Networks (GANs), a deep learning method. When considering the generative process of GANs, and more importantly, when they are intended to generate affect, the utilized training image datasets are of crucial importance.

The experiment presented in this paper, *Forging Emotions*, explores the use of visual emotion datasets and the working processes of GANs for visual affect generation, i.e., to generate images that can convey or trigger specified emotions. It appears that researchers attempting to build visual emotion datasets have almost reached a consensus on the most appropriate methodology for that purpose. Their methodology consists of querying social media for tags with words for emotions (e.g., anger, disgust, fear, sadness) to collect images. A validation procedure is then performed where quite a large number of people confirm if the emotion tags are correct. *Forging Emotions* aims to confront the idea that querying social media with a single word can result in a set of images that truly visualize what an emotion is. Moreover, it aims to question the validation procedure utilized to account for the subjectivity of emotions. Finally, it aims to explore the working processes of GANs for visual affect generation and which visual aspects of the generated images are able to convey or trigger specified emotions.

Background

Affective Computing

Research in affective computing is mainly focused on building systems that can respond and adapt to human emotions [4]. Such systems first need to be able to detect human emotions. For this reason, a great variety of sensors have been developed that can monitor physiological signs that are later utilized for recognizing emotions. The modalities of facial expressions, body gestures, and speech have also been used for recognizing human emotions.

When detecting emotions, it is most common that Paul Ekman's basic emotion theory [5] is utilized, suggesting that all emotions can be derived from a limited set of universal and innate basic emotions. Therefore, the following six basic emotions are considered: anger, disgust, fear, happiness, sadness, and surprise. It is also common that another seventh neutral emotional state is added. Additionally, there are cases where a few other emotions are considered, such as anxiety, excitement, calm, and lust, among others. More recently, two-dimensional models of emotion have gained interest. One very commonly used two-dimensional emotion model is the valence and arousal model [6]. This model proposes that all affective states can be understood as varying degrees of valence (a pleasure – displeasure continuum) and arousal (activation – deactivation).

To a large extent, affective computing research focuses on detecting rather than on generating or triggering emotions. A considerable body of research on affect generation involves synthesizing human-like expressions of emotion through facial features, speech, and gestures for virtual characters and robots [4]. Affective games are a sub-field of affective computing that studies how the gameplay can be adapted according to players' emotions and how emotions can be triggered [7]. Nevertheless, affective games currently adapt to or trigger emotions by design rather than with computational methods.

In conclusion, research on generating affective triggers is highly unexplored. Accordingly, interactive media artworks that utilize affective computing methods also mainly consider affect detection rather than affect generation.

For example, the work Chameleon (2008/2010) by Tina Gonsalves recognized the audience's emotional state, using a facial expression detection system, and classified it to one of the six basic emotions as defined by Ekman [5]. Then, the detected emotions were sent to the video engine that triggered videos from the new corpus Gonsalves created to empathize with the audience's emotional state [8]. Gonsalves created her own facial emotion dataset that was "*more dynamic and aesthetic, engaging and emotionally probing*." Gonsalves' experience working with affective computing methods and visual emotion datasets led her to the conclusion that emotions expressed and monitored in laboratories for scientific research "*don't often correlate to the emotions that form the fabric of our everyday lives*." She also notes: "*There seem to be limited emotions being explored, visually underwhelming databases being used, and the non-ecological settings such as the lab to test responses […] using small groups of subjects with narrow representation, what does the knowledge that science is building about emotions actually mean?*":

Deep Learning Methods

Computational Creativity is a field that has lately regained much interest with the advent of Generative Adversarial Networks (GANs) [9], a deep learning method. If the problem of affect generation is to be considered, then the processes involved in GANs have to be explored. In music, there is the field of affective algorithmic composition (AAC) [10], where an intended affective response always informs the algorithm for music composition. There are many successful applications of GANs within the field of AAC. However, visual affect generation, that is, to generate new images that intend to convey or trigger specific emotions, is still unexplored.

GANs have already been explored by artists, for example, Anna Ridler [11] and Mario Klingemann [12], but not for visual affect generation. One of very few works in the literature that considers the generation of artworks that convey a specifically defined emotion is by David Alvarez-Melis and Judith Amores [13]. In this work, a GAN was trained on 13,000 images of modern art paintings that were emotion-labeled. Alvarez-Melis and Amores concluded that their approach was able to generate artworks with high-level emotional features that agree with psychology literature; for example, red for anger, dark colors for sadness and fear, as well as images that resemble natural landscapes for joy.

GANs have already exhibited the ability to generate images that are hard to distinguish from "real" ones. One of the most well-known GANs is the one developed by NVIDIA to generate new face images that are hard to identify as images not of real people [14]. Additionally, approaches for synthesizing facial affect have also been presented in the literature [15].

However, in most cases, GANs have successfully generated images of objects or concepts that are quite specific in form and structure, e.g., human faces, dogs, landscapes, and others. As a result, the utilized training datasets are comprised of highly similar images. For example, the NVIDIA training dataset includes only photographs of human faces that are also automatically aligned and cropped.

Nevertheless, when higher-level concepts are considered, such as the whole notion of emotions, without being limited to facial expressions, it still remains a question of how to create a suitable training image dataset that could be later used with GANs to generate new affective images, i.e., images that convey or trigger a specified emotion.

Visual Emotion Datasets

Most datasets used for visual emotion analysis are targeted towards applications for affect detection. Even if the problem of visual affect generation is considered, a similar visual emotion dataset is also required for the system to learn how to generate affective images. For that purpose, in the following, the most commonly used visual emotion datasets are summarized. It should be noted that this discussion will not include datasets of facial expressions that are beyond the scope of the presented experiment *Forging Emotions*.

Emotion researchers have constructed quite a few visual emotion datasets. They commonly define specific categories of images according to their research questions and then hand-select images. The International Affective Picture System (IAPS) was designed to provide a standardized set of pictures for studying emotion and attention and was developed by the National Institute of Mental Health Center for Emotion and Attention at the University of Florida [16]. IAPS is the most widely used visual emotion dataset in hundreds of behavioral and neuroimaging studies. It includes over 1,000 images depicting people experiencing various emotions (e.g., sad, fearful, angry), erotic couples, funerals, dirty toilets, cityscapes, landscapes, wars and disasters, mutilated bodies, baby animals, and many more. Other visual emotion datasets constructed by emotion researchers are the Geneva Affective Picture Database (GAPED) [17] with 730 images and the Nencki Affective Picture System (NAPS) [18] with 1,256 images. In the GAPED dataset, images of spiders, snakes, and scenes of moral or legal violations were selected for negative images, whereas for positive images, mainly humans, animal babies, and landscapes were used. Images in the NAPS dataset belong to five broad categories: people, faces, animals, objects, and landscapes. The "object" category is a very broad class in which a wide range of clearly visible objects, foods, or vehicles were depicted without humans or animals present. Pictures in all categories included stimuli for different emotions. For example, images in the "people" category included alive, injured, or dead human bodies. In all the above-mentioned visual emotion datasets, images were evaluated on valence and arousal.

Computer scientists also have built visual emotion datasets. Mikels et al. [19] created the Subset A of IAPS (IAPSa) by selecting 395 pictures from IAPS and then categorizing them into eight discreet categories (anger, disgust, fear, sadness, amusement, awe, contentment, and excitement) by conducting a user study. Machajdik and Hanbury [20] created another dataset of 807 artistic photographs downloaded from DeviantArt, an online social community for artists and art enthusiasts. The same eight basic emotion categories were also used in this dataset. The associated emotion for each image was the label given by its owner on DeviantArt. A more recent dataset is the Open Affective Standardized Image Set (OASIS) [21]. It contains 900 images, collected using the Google Images search engine, depicting a broad spectrum of themes, including humans, animals, objects, and scenes. Each image was rated on valence and arousal by recruiting participants through Amazon's Mechanical Turk (MTurk).

You et al. [22] aimed at creating a large visual emotion dataset and remark that all the previously mentioned datasets are significantly smaller than those used for other computer vision tasks. For example, the ImageNet dataset [23] used for object recognition as well as image generation with GANs contains more than 14 million images. Hence, You et al. collected over 3 million images from Flickr and Instagram labeled with one of the eight basic emotions. Furthermore, they also employed MTurk to verify the emotion labels associated with images. Finally, the created dataset contains in total over 23,000 images with verified emotion labels, which currently is the largest visual emotion dataset available.

This discussion on visual emotion datasets can conclude that computer science researchers generally turn into online resources, and more often, to social media, to collect images associated with emotions with single-word queries. Then, most commonly, MTurk is employed so that many different people validate the emotion conveyed or triggered by each image. Thus, it seems that researchers in the field have almost reached a consensus on the most appropriate methodology, and what remains to be completed is to build a vast visual emotion dataset that would contain millions of images.

Forging Emotions

It is uncertain that the previously described visual emotion datasets can genuinely visualize how emotions are experienced or evoked. An emotion is not just a word, such as the ones used to query social media, but a network of associations. According to psychology research, each individual relates an emotional percept or event in many different ways to a multitude of past emotional experiences [24]. The basic idea is that, for example, the emotion of happiness is not just a word but a concept that can be represented by a network with nodes of past emotional experiences. Thus, the network of happiness could include nodes for the emotional experience of a child being born, a vacation, being with friends and family, and many more. Everyone experiences emotions differently, and thus emotional networks of associations are individual, i.e., each individual may have different emotional experiences related to happiness. However, according to psychology literature, emotional networks share a common basic structure for most people [24].

Forging Emotions aims to confront the idea that querying social media with a single word can result in images that genuinely visualize how emotions are experienced or evoked. Moreover, it seeks to question the process of employing MTurk to account for the subjectivity of emotions. For that purpose, two new image datasets were created by querying Instagram with the hashtags #sad and #happy. The collected images were not verified in any way.

All the previously mentioned visual emotion datasets do not consider the emotion of happiness, although it was included in the six basic emotions originally introduced by Ekman. In social media behavior studies, it is often concluded that one always tries to appear attractive, happy, and clever [25]. In the context of the *Forging Emotions* experiment, it seems like a paradox not to explore the happy emotion while social media are utilized to understand emotions. The #happy dataset poses the question of whether this is the emotion that can be better understood from social media data or if shared emotional experiences on social media are authentic.

This experiment also aims to test whether new images can be composed that visualize and evoke the emotions of happiness and sadness, or in some manner can "forge emotions", which relates to commonly naming the generator component of a GAN a "forger." Thus, the experiment was named *Forging Emotions*.

GANs have been successfully applied for generating new images of human faces, dogs, landscapes, and many more. However, the criticism that follows the use of GANs is that they are not interpretable, and they are often considered black boxes. Consequently, interpreting and understanding what a GAN has learned is an active research topic within Explainable Artificial Intelligence (XAI) [26], which has gained much interest lately. In the recent work of Bau et al. [27], the aim was to understand if a GAN learns composition or if it purely memorizes pixel patterns. The conclusion was that GANs indeed learn aspects of composition and that certain neuron units have learned specific features of the taught domain.

Another exploration strategy for interpreting and understanding deep learning models is the black-box exploration. In this type of exploration, only the training dataset and the deep learning model's output are used to examine its behavior and provide insight into its interpretations [28]. This approach is similar to the one Ridler [11] describes in which she used GAN images as a mirror to her own drawing process.

In this respect, *Forging Emotions* will also perform a black-box exploration to understand what a GAN has learned when trained with a visual emotion dataset.

Experiments

The experiments were performed with the implementation, and the proposed architecture guidelines, for stable Deep Convolutional GANs (DCGANs) [29]. The training images were not pre-processed, except scaling them to the range of the tanh activation function [-1, 1]. All models were trained with mini-batch stochastic gradient descent (SGD) with a mini-batch size of 64 for 100,000 epochs. All weights were initialized from a zero-centered normal distribution with a standard deviation of 0.02. In the LeakyReLU, the slope of the leak was set to 0.2 in all models. The Adam optimizer was used with a learning rate of 0.0002. Finally, the momentum term β_1 was set to 0.5. The images were generated at a 64x64 pixel resolution. In all the following experiments, the DCGANs were trained with the same parameters and for the same number of epochs.

Forging Sadness

Initially, the image dataset created for the *Forging Emotions* experiment with 30,000 Instagram images that include the hashtag #sad was used to train the DCGAN. Then, new images were generated and are shown in Figure 1. The same procedure was followed for the dataset created by You et al. [22] for the emotion of sadness, which included 2,635 images. The images generated in this case are shown in Figure 2.

It can be observed that all the generated images shown in both figures are, in some way, reminiscent of abstract paintings. In both cases, some images would seem like abstract human figures or, in other cases, abstract landscapes or cityscapes. In Figures 1 and 2, it can be noticed that there are more images with dark colors, which is in line with psychology literature that generally associates dark colors with unpleasant emotions. A distinct differentiation is that Figure 1 includes more images that look like text which is due to the different types of images included in the two utilized training datasets.

The dataset utilized for generating the images shown in Figure 2 was validated by humans and mainly consisted of images depicting humans with facial or body expressions of sadness. The dataset for sadness created by You et al. also included quite a few images of pets. Other types of images constitute only a tiny proportion of the dataset and include text images, graves or graveyards, dead flowers or plants, winter landscapes, etc. The You et al. dataset [22] included only 2,635 images for the emotion of sadness, and it is evident that a significantly larger dataset is required for the GAN to be able to generate less abstract images.

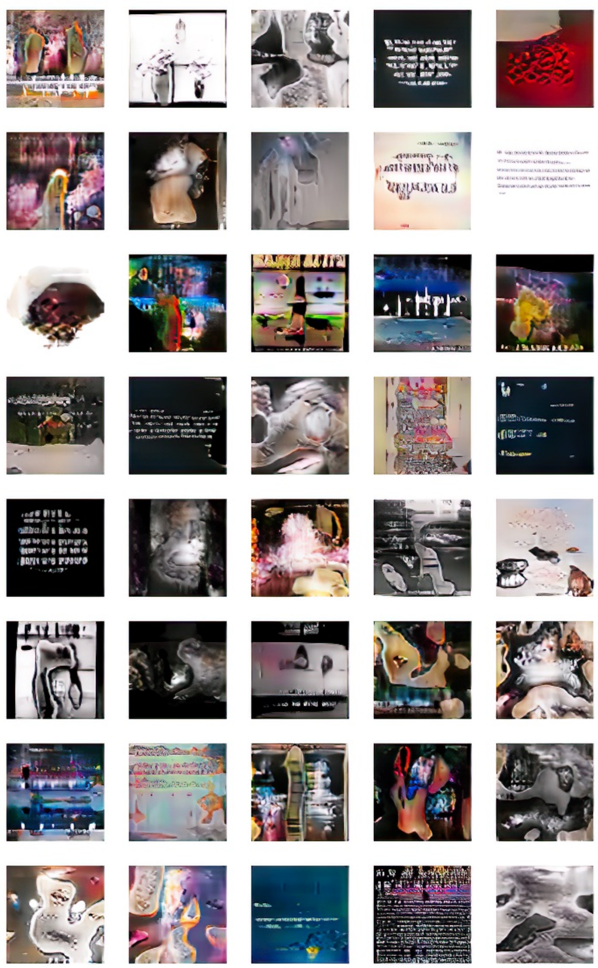

Figure 1: The DCGAN generated images when trained with 30,000 Instagram images with the hashtag #sad. (Full-size image available at https://pasteboard.co/JKcqhxo.png)

On the other hand, the image dataset collected for the *Forging Emotions* experiment included predominately text images. Our dataset also included many images depicting humans, pets, dead plants, and winter landscapes. Additionally, it included images of comics, movie stills, bad food, and other types of images that were not present in the dataset by You et al. [22]. Finally, since the sadness dataset created for the *Forging Emotions* experiment was not validated, it included some images that were not associated with the emotion of sadness but rather with the meaning of the word sad for characterizing something as inadequate or unfashionable.

Nevertheless, it becomes evident that collecting social media images with single-word queries cannot guarantee that the whole range of how people experience emotions can be included. For example, we cannot find any images for sadness in either of the two considered datasets that depict, for instance, funerals, wars, disasters, or mutilated bodies, as was the case in the IAPS dataset created by emotion researchers. Moreover, it is most probable that such images would be removed, even if they were posted on social media, due to their policies in order.

In conclusion, although the employed validation method with MTurk ensures that some irrelevant pictures would be removed from the dataset, it cannot accommodate for the fact that several aspects of how sadness is experienced are missing. Social media users' behavior, along with emotion research, should be further studied to identify multiple image sources and the emotional networks of associations that constitute how emotions are experienced. Finally, significantly larger datasets should be built for training GANs.

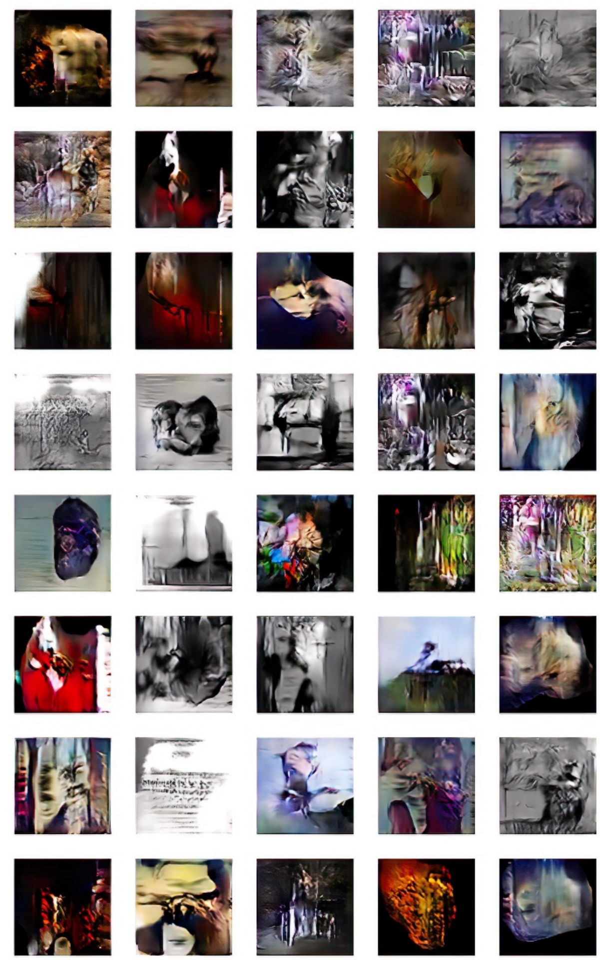

Figure 2: The DCGAN generated images when trained with the image dataset created by You et al. for sadness. (Full-size image available at https://pasteboard.co/JKcpdeQ.png)

Forging Happiness

Another DCGAN was trained using the image dataset created by collecting 30,000 Instagram images that include the hashtag #happy, and the generated images are shown in Figure 3.

Again, all generated images are reminiscent of abstract paintings, including abstract human figures, landscapes, and cityscapes. Images in Figure 3 have noticeably brighter colors than those in Figures 1 and 2, which is also in line with psychology literature that generally associates bright colors with pleasant emotions.

Instagram users post a greater variety of images when they use the hashtag #happy. They also post many pictures depicting humans but similarly many images with luscious food, photos from vacations that include landscapes and cityscapes, images showing an object of desire, babies, and many more.

Even though the whole range of how people experience happiness is still not included in the collected images, what becomes more evident in this case is that it is questionable that the generated images can visualize or trigger the emotion of happiness. Generating images that merely correspond to basic psychology findings, e.g., bright colors like in the work of Alvarez-Melis and Amores [13], does not seem adequate for visual affect generation. For example, although the red color is generally associated with anger in psychology findings, this is not always the case. In an image depicting a natural landscape at sunset, the sky is often red, and it would trigger positive emotions for many people.

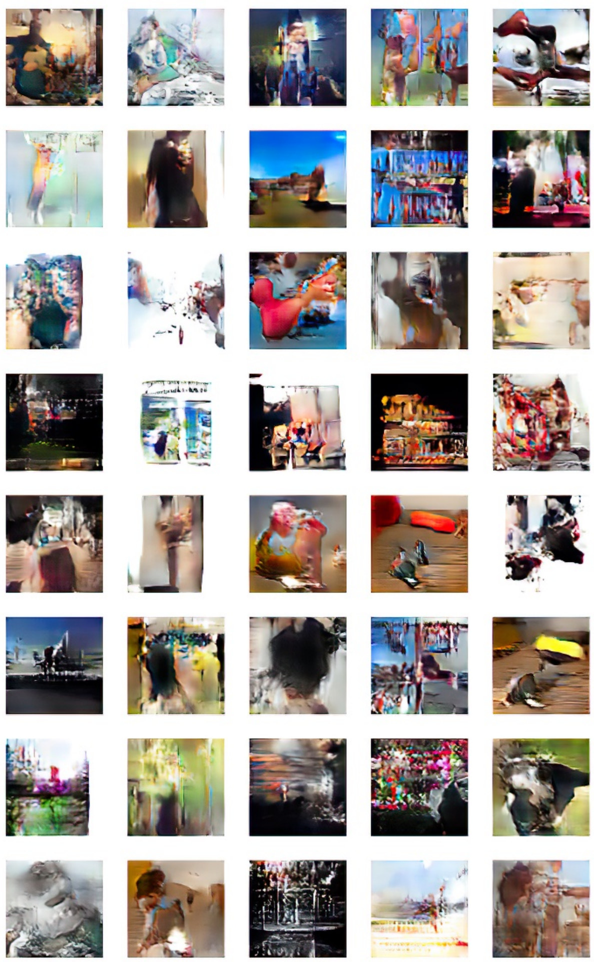

Figure 3: The DCGAN generated images when trained with 30,000 Instagram images with the hashtag #happy. (Full-size image available at https://pasteboard.co/JKcrJO2.png**)**

The two first images on the second row of Figure 3 quite clearly depict women posing. These images were generated because the #happy dataset included many images posted by women posing with their new clothes or in other cases posting selfie images to state that they are happy for various reasons. Accordingly, the third image on the second row of Figure 3 depicts a natural landscape like the ones often posted on Instagram from vacations. However, as exhibited in the case of the *Forging Sadness* experiment, significantly larger datasets should be built for training GANs that include more aspects of how happiness is experienced.

More importantly, recent research efforts on understanding the structure of trained GAN could shed light on the different features learned from visual emotion datasets. This knowledge could be applied towards training compositional GANs that could create genuinely novel compositions and combine different elements from different types of images. Thus, the subject matter of generated images would be better directed for visually generating affect.

Conclusions

The experiment *Forging Emotions* was doomed to fail in its initial set goal to generate new images that convey or trigger the emotions of happiness and sadness. Nevertheless, its execution allowed us to draw two main conclusions and instigate research on the identified issues. First, the methodology used so far by researchers to build datasets for describing high-level concepts such as emotions is not sufficient. This conclusion is also in line with the comments artists made while working on affective computing methods; for example, the issue raised by Gonsalves on whether the datasets used in scientific research, especially in the case of emotions, can really "*correlate to the emotions that form the fabric of our everyday lives*," remains unresolved. It seems that the collected images based on single-word queries can indeed generally be characterized as related to an emotion such as happy or sad, but not able to describe the full range of the emotional experience of happiness or sadness. It would appear as a more appropriate methodology to first explore a network of associations for an emotion that describes the emotional experience rather than just the name of an emotion. Additionally, a greater variety of online sources should be considered rather than only social media. Finally, the contribution of artists in the creation of visual emotion datasets would be most valuable.

The last conclusion of this experiment is that for visual affect generation, research efforts should be targeted towards better understanding the structure of trained GANs and compositional GANs. GANs, in many cases, are able to understand the different elements used to compose, for example, a scene. If visual emotion datasets are built according to emotional networks of associations, they will comprehensively visualize how each emotion is experienced. Then, methods and frameworks for understanding and interpreting GANs will shed light on the different features learned from visual emotion datasets. This knowledge could be applied to train GANs that can create genuinely novel compositions and combine different elements from different types of images. Thus, not only the colors but also the subject matter of generated images will be better directed for visually generating affect.

Res. Methods, vol. 46, no. 2, pp. 596–610, Aug. 2014.

References

[1] K. R. Scherer, “What are emotions? And how can they be measured?,” *Soc. Sci. Inf.*, vol. 44, no. 4, pp. 695–729, Dec. 2005.

[2] J. Robinson, *Deeper than reason : emotion and its role in literature, music, and art*. Clarendon Press, 2005.

[3] K. Kwastek, *Aesthetics of interaction in digital art*. MIT Press, 2013.

[4] R. A. Calvo, S. D’Mello, J. M. Gratch, and A. Kappas, Eds., *The Oxford handbook of affective computing*. Oxford University Press, 2015.

[5] P. Ekman, “Basic Emotions,” in *Handbook of Cognition and Emotion*, Chichester, UK: John Wiley & Sons, Ltd, 2005, pp. 45–60.

[6] T. Eerola and J. K. Vuoskoski, “A comparison of the discrete and dimensional models of emotion in music,” *Psychol. Music*, vol. 39, no. 1, pp. 18–49, Jan. 2011.

[7] R. Lara-Cabrera and D. Camacho, “A taxonomy and state of the art revision on affective games,” *Futur. Gener. Comput. Syst.*, vol. 92, pp. 516–525, Mar. 2019.

[8] T. Gonsalves, N. Berthouze, and M. Iacobini, “The Chameleon Project,” *Second Nat.*, no. 3, pp. 138–163, 2010.

[9] I. Goodfellow *et al.*, “Generative Adversarial Nets,” in *NIPS*, 2014, pp. 2672–2680.

[10] D. Williams, A. Kirke, E. R. Miranda, E. Roesch, I. Daly, and S. Nasuto, “Investigating affect in algorithmic composition systems,” *Psychol. Music*, vol. 43, no. 6, pp. 831–854, Nov. 2015.

[11] A. Ridler, “Repeating and mistranslating: the associations of GANs in an art context,” in *Machine Learning for Creativity and Design, NIPS 2017 Workshop*, 2017.

[12] M. Klingemann, “Quasimondo | Mario Klingemann, Artist working with Code, AI and Data,” 2019. [Online]. Available: http://quasimondo.com/. [Accessed: 15-Dec-2019].

[13] J. Alvarez-Melis, David and Amores, “The Emotional GAN: Priming AdversarialGeneration of Art with Emotion,” in *Machine Learning for Creativity and Design, NIPS 2017 Workshop*, 2017.

[14] T. Karras, S. Laine, and T. Aila, “A Style-Based Generator Architecture for Generative Adversarial Networks,” in *IEEE Conference on Computer Vision and Pattern Recognition*, 2019, pp. 4401–4410.

[15] M. Park, H. G. Kim, and Y. M. Ro, “Photo-realistic facial emotion synthesis using multi-level critic networks with multi-level generative model,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2019, vol. 11296 LNCS, pp. 3–15.

[16] P. J. Lang, M. M. Bradley, and B. N. Cuthbert, *International affective picture system (IAPS)*. NIMH, Center for the Study of Emotion & Attention, 2005.

[17] E. S. Dan-Glauser and K. R. Scherer, “The Geneva affective picture database (GAPED): A new 730-picture database focusing on valence and normative significance,” *Behav. Res. Methods*, vol. 43, no. 2, pp. 468–477, Jun. 2011.

[18] A. Marchewka, Ł. Żurawski, K. Jednoróg, and A. Grabowska, “The Nencki Affective Picture System (NAPS): Introduction to a novel, standardized, wide-range, high-quality, realistic picture database,” *Behav.*

[19] J. A. Mikels, B. L. Fredrickson, G. R. Larkin, C. M. Lindberg, S. J. Maglio, and P. A. Reuter-Lorenz, “Emotional category data on images from the international affective picture system,” *Behav. Res. Methods*, vol. 37, no. 4, pp. 626–630, Nov. 2005.

[20] J. Machajdik and A. Hanbury, “Affective image classification using features inspired by psychology and art theory,” in *Proceedings of the international conference on Multimedia - MM ’10*, 2010, p. 83.

[21] B. Kurdi, S. Lozano, and M. R. Banaji, “Introducing the Open Affective Standardized Image Set (OASIS),” *Behav. Res. Methods*, vol. 49, no. 2, pp. 457–470, Apr. 2017.

[22] Q. You, J. Luo, H. Jin, and J. Yang, “Building a Large Scale Dataset for Image Emotion Recognition: The Fine Print and The Benchmark,” *Thirtieth AAAI Conf. Artif. Intell.*, Feb. 2016.

[23] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, “ImageNet: A Large-Scale Hierarchical Image Database,” in *CVPR09*, 2009.

[24] J.-M. Fellous and J. Robinson, “A Mechanistic View of the Expression and Experience of Emotion in the Arts,” *Am. J. Psychol.*, vol. 119, no. 4, p. 668, Dec. 2006.

[25] D. Freitas, *The happiness effect : how social media is driving a generation to appear perfect at any cost*. Oxford University Press, 2017.

[26] F. Xu, H. Uszkoreit, Y. Du, W. Fan, D. Zhao, and J. Zhu, “Explainable AI: A Brief Survey on History, Research Areas, Approaches and Challenges,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2019, vol. 11839 LNAI, pp. 563–574.

[27] D. Bau *et al.*, “GAN Dissection: Visualizing and Understanding Generative Adversarial Networks,” *arXiv Prepr. arXiv1811.10597*, Nov. 2018.

[28] J. Wang, “Interpreting and Diagnosing Deep Learning Models: A Visual Analytics Approach,” The Ohio State University, 2019.

[29] A. Radford, L. Metz, and S. Chintala, “Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks,” *arXiv Prepr. arXiv1511.06434*, Nov. 2015.

Author Biography

Amalia Foka (b. Greece) is a creative Artificial Intelligence researcher and educator who explores the intersection of computer science and art. Her work makes use of different artificial intelligence technologies and social media data to generate and study art. Her work has been presented and published internationally including the Leonardo Journal (MIT Press); WRO Media Art Biennale; EVA London and many more. She is currently an Assistant Professor in Computer Science: Multimedia Applications for the Arts at the School of Fine Arts, University of Ioannina, Greece, where she teaches courses on creative coding, creative interactive systems, and generative AI art. She holds a BEng in Computer Systems Engineering (1998) and an MSc in Advanced Control (1999) from the University of Manchester Institute of Science & Technology (UMIST) in the United Kingdom. She also holds a Ph.D. in Robotics (2005) from the Department of Computer Science of the University of Crete.